Parker

BNL-19389

401320

The Frequency of "Rare" Protein Variants in Marshall Islanders and other Micronesians[1]

by

James V. Neel[2]

Robert E. Ferrell[2]

Robert A. Conard[3]

1. Research supported by the Atomic Energy Commission

2. Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48104

3. Brookhaven National Laboratory, Associated Universities, Inc., Upton, L. I., N. Y. 11973.

Recently, in connection with an ongoing medical study of certain Marshall Islanders prompted by the Bikini incident (1), we have had the opportunity to examine a series of blood samples from these Islanders for the occurrence of both polymorphisms and rare variants of serum proteins and erythrocyte enzymes. The results, when treated in the usual fashion and combined with the findings of others on Micronesians, there may be suggest/a lower frequency of rare variants in this group than in such groups as Japanese, Caucasians, or South American Indians. There are some bothersome problems in comparisons across groups sampled in different ways, however, which make the usual statistical contrasts impossible; some of these problems are aired.

THE POPULATION

The study population is composed of persons now residing on Ebeye, Rongelap, and Majuro Islands, for the most part related to one another as members of nuclear families. The number of independent genomes in the sample is thus considerably less than the number of persons. Approximately half of the children in the sample were born to parents inadvertently radiated as a result of fall-out from a nuclear explosion at the time of the Bikini test, in 1954. However, as will be apparent under RESULTS, the question of a radiation effect will not arise in any substantial manner.

METHODS

All samples were collected in 12 ml vacutainers (Becton-Dickinson) with an ACD anticoagulant. The samples were shipped by air, on ice, from Kwajalein Atoll, Marshall Islands to Honolulu, Hawaii for transshipment to Ann Arbor. Washed red blood cells and plasma were stored at -70° C prior to typing.

The conditions for electrophoresis and typing of systems 1, 2, 8, 9, 10, 12, 13, 15, 16, 18 and 20-24 in Table 1 were carried out as described previously (2). Electrophoresis of systems 14 and 19 and staining of 14 employed the method of Spencer, Hopkinson and Harris (3), and staining for system 19 employed the positive staining method as reported by Peters, Hopkinson and Harris (4). System 3 was determined by the method of Charlesworth (5), system 4 by the method of Tashian (6), system 5 by the method of Chen, Anderson and Giblett (7), system 6 by the method of Weitkamp (8), system 11 by the method of Edwards, Hopkinson and Harris (9), and system 20 also by the methods of Weitkamp et al (10) and Tanis et al (11).

FINDINGS

1. The polymorphisms.--Genetic polymorphisms were observed in six of the systems: haptoglobin, phosphoglucomutase-1, adenosine deaminase, acid phosphatase, 6-phosphoglucose dehydrogenase, and group specific component. Phenotype and allele frequencies are given in Table 2.

Varying numbers of determinations for the systems listed in Tables 1 and 2 are due to a considerable time lapse between collection of the samples and their receipt in the laboratory. As a consequence, either no pattern or an unclear pattern (even after repetition) was occasionally obtained for a given trait. Three persons had the PGM_1-7 phenotype. Because of the occurrence of this variant in polymorphic proportions elsewhere in Micronesia (see below), we tabulate it with the polymorphisms. These three individuals were related as siblings, one of whose parents was tested and normal, the other untested.

2. Rare variants.--One rare variant was observed in a total of 4,047 determinations. This was a fast albumin variant detected by only one of the three screening systems in use for albumins, namely, the pH 5.0 sodium acetate buffer system of Weitkamp et al (10). The variant is illustrated in Fig. 1. Its electrophoretic behavior appears similar to that of albumin Medón as reported by Weitkamp et al (10). It occurred in a girl, aged 5, whose mother was normal (and not exposed to fall-out) and father not available for study. The fact that the variant has been demonstrated in only a single person with only one technique leaves its identification somewhat unsatisfactory, but the situation cannot be improved upon at present.

DISCUSSION

1. The polymorphisms.--Gene frequencies for the 6 polymorphisms encountered all fell within the rather considerable range reported in other studies of Micronesians (12, 13, 14, 15, 16, 17, 18, 19).

2. The rare variants.--In Table 2 we have summarized not only our own findings but also the results of all the other studies of rare variants of 24 systems in Micronesians which we have been able to locate in the literature. Any effort to treat "rare" variants involves some arbitrary decisions; no approach is apt to find universal acceptance at this time. We exclude from this summary any variant which for the totality of the representatives of the population studied to date, occurs in more than 2.0% of the group (one of the conventional definitions of a polymorphism). By this definition of rare variant, we exclude from the tabulation the polymorphisms involving types 3 and 7 of the PGM_1 system reported by Blake et al (12) in the Western Caroline Islands, and the polymorphism for the type 2 of phosphoglycerate kinase reported by these same authors. This same convention will require us to eliminate from the summary of rare variants in Amerindians (see below) the Yanomama-2 variant of albumin, a variant thus far encountered in a single tribe, but there with a gene frequency of 0.08. In sampling populations where rather close biological relationships between individuals can scarcely be avoided, it is well to remember that even in samples of 1000, a variant limited to members of a single extended kindred may assume the proportions of a polymorphism as here defined. In the case of the variants excluded from the Micronesia count, a variant with the electrophoretic mobility of type 7 PGM_1 has been encountered at several localities in the Far East, and it seems likely this is a "widespread" polymorphism. However, the same cannot be said for the variant with the mobility of PGM_1-3 or the variant with the mobility of PGK-2. They may be "private" polymorphisms. If further studies in Micronesia revealed these alleles

to have so restricted a distribution that their frequency in the entire sample dropped below 2 percent, then they would of course enter the list of rare variants.

The final column of Table 2 presents the frequency per 1000 of variants for each system, and the unweighted average of all the systems. An unweighted average is employed to prevent an extensive study of a single system from dominating the picture. By this approach, the average frequency with which variants are encountered in the systems under consideration, on the basis of 16,724 determinations, is 0.9/1000 observations.

Some of the Micronesian Islands have had extensive contact with other ethnic groups, and there has in addition been active exchange among the various islands (22,23). The possibility must always be borne in mind that a rare variant may have been introduced from the outside, as pointed out by Blake et al (12) in connection with the LDH variant they encountered on Faraulep Atoll, which is similar to the Calcutta-1 variant widespread in India. We note in this connection that the CA I variant of Tashian et al (20) encountered in the Chamorros of Guam and Saipan is electrophoretically similar to a variant encountered in Filipinos (24), Indonesians (25), and Japanese (26); Guam and Saipan have well-documented historical contacts with these areas. On the other hand, we restate the well-recognized caveat that electrophoretic identity of two variants is not synonymous with biochemical identity.

There is rapidly accumulating an extensive literature on the frequency of rare variants in a variety of populations. We mentioned earlier the

problems in comparing populations sampled differently. In addition, the comparison of variant frequencies from different laboratories, and especially average frequencies when the mix of systems studied may vary widely, is to be approached with caution. For the purposes of this preliminary comparison, we will content ourselves with reference to two series which originate in this laboratory and involve essentially the same proteins and techniques as the original material herein reported, plus a third composite series for Caucasians. We will again use an unweighted average, based on as many of the systems listed in Table 2 as are covered in the references cited. By this convention, the variant frequency for South American Indians, based on systems 1, 2, 8-10, 11-16, 18 and 20-24 of Table 2, is 1.7/1000 (2, and unpublished). The lower frequency than in our last publication on this subject (2) is due to the exclusion of the Yanomama-2 albumin variant from this calculation because its frequency in the total sample now exceeds 2 percent (11). The frequency in West European Caucasians, based on systems 1, 2, 8-16, 18-24 and the data of Fleischer and Mohr (27); Moullec et al (28); Sick et al (29); Bajatzadeh and Walter (30); Fine (31); Rex-Kiss and Fesus (32); and Harris, Hopkinson, and Robson (33) is 2.7/1000, and for Japanese adults, based on systems 1, 2, 4, 8-10, 12, 13, 15, 16, 18, and 20-24, 1.9/1000 (34).

As noted, there are obvious difficulties in statistical comparisons between these series. Our own Micronesian data, and we suspect to some extent that of others, contains closely related individuals, as does our data on Amerindians. By contrast, the data on Caucasians and Japanese

presumably contains very few biologically related persons. Elsewhere we have argued that it is appropriate to compare the average number of variants per system per thousand determinations in populations ascertained in these different fashions, but not the number of different variants (2), and we maintain that position here. The logic is that a rather considerable shuffling of tribal populations, such as occurred in the detribalization of the ancestors of modern Japanese or Europeans, should not alter the total number of variants present. However, conventional statistical contrasts of total frequencies in these various populations seem inappropriate. In particular, estimates of frequencies from studies of populations such as Micronesians or Amerindians are quite susceptible to a "jackpot" effect--one island with a high frequency variant could markedly alter the picture. However, some 27 islands have been sampled to date, comprising a rather representative group.

Taken at face value, there is a 3-fold range among ethnic groups in the frequency of rare variants as defined, these variants occurring in Micronesians with approximately half the frequency in which they have been encountered in several other ethnic groups. However, this apparent difference hinges, in part, on the definition of rare variant, which in turn is intimately related to the size and nature of the sample. Thus, if as the study of Micronesians is extended it becomes clear that the PGM_1-3-like variant encountered in 49 persons (12) and the PGK-2-like variant encountered in 39 persons (12) are sharply localized and no or few additional examples of the variant are encountered, then either or both of these might drop below the arbitrary 2 percent frequency level and have to be classed as rare variants, with a marked impact on the average

frequency of such variants. This possibility brings out clearly the sensitivity to the vagaries of the sampling process of any approach which determines the frequency of rare variants from a semi-exhaustive sampling of discrete demes in a minimally disturbed population. Otherwise stated, the estimate must have a large variance.

The properties of the samples of Micronesians and Amerindians might be rendered more comparable to those of Caucasians and Japanese by eliminating the old and the young from the former two samples, but the basic issue would still remain: the individuals in the former two samples will be more related than those in the latter two. The most practical way to meet this issue seems an extension of the sample, thus reducing the impact of any one "jackpot". In addition, it is suggested that the definitive treatment of comparative variant frequencies which a larger sample will permit must consider a variety of definitions of rare variant and/or analyze the total heterozygosity of the population.

The problem is not a trivial one. The frequency of such variants in natural populations is maintained by a complex balance between selection, mutation, and population structure. Basic parameters though they be, both the manner of action of selection and the rate and types of mutation in higher organisms remain poorly understood. Both selection and mutation may be studied directly, i.e., thru surveys followed by detailed family studies of each variant, or indirectly, i.e., through the manipulation of population parameters (35,36). The former approach is much more laborious than the latter, and it is tempting to pursue the easier course, but the latter approach is only as sound as the estimates of variant frequency (which should be based on as total a

population study as possible) and an understanding of the structure of the populations on which the estimates are to be based. This is, of course, true for plant and animal as well as human populations. In the present instance, we are of course intrigued by the possibility that the mix of these factors differs in Micronesians from the other groups cited, e.g., lower mutation rates or greater stochastic loss of new mutants in the Micronesians. Before these possibilities can be profitably pursued by the indirect approach, sample size must be adequate. There is clearly some minimal sample necessary to a trustworthy estimate of the frequency of rare variants in populations like Micronesians and Amerindians, and most investigators would probably agree that the Micronesians do not yet approach that minimum. Unfortunately, our knowledge of the clustering of specific rare variants in relatively undisturbed populations is still so scanty that further experience is necessary prior to setting that reasonable minimum.

SUMMARY

Blood specimens from a sample of 187 Marshall Islanders were studied with reference to variants of 22 serum proteins and erythrocyte enzymes. Six of the traits studied exhibited genetic polymorphisms (adenosine deaminase, phosphoglucomutase$_1$, acid phosphatase, 6-phosphogluconate dehydrogenase, haptoglobin, group specific component). There was in addition one "rare" variant (of albumin) in 4,047 determinations. These results on rare variants have been combined with those of others on Micronesians and the frequency of rare variants in Micronesians compared with the frequencies in West

European Caucasians, Japanese, and Amerindians. There are many difficulties in such comparisons, and although the observed values for the four ethnic groups differ by a factor of three, the Micronesians exhibiting the lowest frequency, it is felt that no conclusions concerning differences between ethnic groups can be drawn at this time.

REFERENCES

(1) Conard, R. A., et al 1975 Twenty year report of medical findings on the people of Rongelap and Utirik Islands accidentally exposed to fall-out in 1954. Brookhaven National Laboratory Report, No. 50424.

(2) Tanis, R. J., J. V. Neel, H. Dovey, and M. Morrow 1973 "The Genetic Structure of a Tribal Population, the Yanomama Indians. IX Gene Frequencies for 18 Serum Protein and Erythrocyte Enzyme Systems in the Yanomama and Five Neighboring Tribes: Nine New Variants". Amer. J. Hum. Genet. 25, 655.

(3) Spencer, N., D. A. Hopkinson, and H. Harris 1964 "Phosphoglucomutase Polymorphism in Man", Nature 204, 742.

(4) Peters, J., D. A. Hopkinson, and H. Harris 1973 "Genetic and non-genetic variation of triose phosphate isomerase isozymes in human tissue", Ann. Hum. Genet., Lond, 36, 297.

(5) Charlesworth, D. 1972 "Starch gel electrophoresis of four enzymes from human red blood cells: glyceraldehyde-3-phosphate dehydrogenase, fructoaldolase, glyoxylase II, and sorbital dehydrogenase", Ann. Hum. Genet., Lond., 35, 297.

(6) Tashian, R. E. 1969 "The Esterases and Carbonic Anhydrases of Human Erythrocytes", in Biochemical Methods in Red Cell Genetics, J. J. Yunis, Ed., Academic Press, p 307.

(7) Chen, S., S. E. Anderson, and E. R. Giblett 1971 "2,3-Diphosphoglycerate mutase: its demonstration by electrophoresis and the detection of a genetic variant", Biochem. Genet., 5, 481.

(8) Weitkamp, L. R. 1973 "Electrophoresis of galactose-1-phosphate

[redacted]

5009638

uridyl transferase", Isozyme Bulletin No. 6, Cornell University, Ithaca, New York.

(9) Edwards, Y. H., D. A. Hopkinson, and H. Harris 1971 ,"Inherited variants of human nucleoside phosphorylase", Ann. Hum. Genet., Lond., 34, 395

(10) Weitkamp, L. R., F. M. Salzano, J. V. Neel, F. Porta, R. A. Geerdink, and A. L. Tarnoky 1973 ,"Human Serum Albumin: twenty-three genetic variants and their population distribution", Ann. Hum. Genet., Lond. 36, 381.

(11) Tanis, R., Ferrell, R. E., Neel, J. V., and Morrow, M. 1974 Albumin Yanomama-2, a 'private' polymorphism of serum albumin. Ann. Hum. Genet., Lond. 38, 179.

(12) Blake, N. M., K. Omoto, R. L. Kirk, and D. C. Gajdusek 1973 "Variation in Red Cell Enzyme Groups among Populations of the Western Caroline Islands, Micronesia", Amer. J. Hum. Genet., 25, 413.

(13) Douglas, R., J. Jacobs, J. Sherliker, and J. M. Staveley 1961 , "Blood Groups, Serum Genetic Factors, and Hemoglobins in Gilbert Islanders", N. Z. Med. J., 60, 146.

(14) Plato, C. C. and M. Cruz 1966 "Blood Groups and Haptoglobin Frequencies of the Trukese of Micronesia", Acta genet. Basel, 16, 74.

(15) Blumberg, B. S. and Z. Gentile 1961 , "Haptoglobin a and Transferrins of Two Tropical Populations", Nature, 189, 897.

(16) Hainline, J., P. Clark, and R. J. Walsh 1969 , "ABO, Rh and MNS Blood Typing Results and other Biochemical Traits in the People of the Yap Islands", Arch. Phys. Anthrop. Oceania, 4, 64.

(17) Morton, N. E. and M. Yanamoto 1973 , "Blood Group and Haptoglobins in the Eastern Carolines", Am. J. Phys. Anthrop., 38, 695.

(18) Plato, C. C. and M. Cruz 1967 , "Blood Group and Haptoglobin Frequencies of the Chamorros of Guam", Am. J. Hum. Genet. 19, 722.

(19) Yamamoto, M. and L. Fu 1973 , "Red Cell Isozymes in the Eastern Carstines", Am. J. Phys. Anthrop., 38, 703.

(20) Tashian, R. E., C. C. Plato, and T. B. Shows 1963 ,"Inherited Variant of Erythrocyte Carbonic Anhydrase in Micronesians from Guam and Saipan", Science, 140, 53.

(21) Tashian, R. E., unpublished results.

(22) Simmons, R. T., A. J. Graydon, D. C. Gajdusek, and P. Brown. 1965 , Blood group genetic variations in natives of the Caroline Islands and in other parts of Micronesia, with a supplement table of voyages affecting Micronesian Islands by S. H. Riesenberg. Oceania 36, 132.

(23) Riesenberg, S. H. 1965 Tables of Voyages Affecting Micronesian Islands. Oceania. 36, 155.

(24) Lie-Injo, L. E. 1967 Red cell carbonic anhydrase Ic in Filipinos. Am. J. Hum. Genet., 19, 130.

(25) Lie-Injo, L. E. and H. G. Pooy Dey. 1970 Phosphoglucomutase, carbonic anhydrase and catalase in Indonesians. Hum. Hered. 20, 215.

(26) Ueda, N. New Japanese variant of human erythrocyte carbonic anhydrase. Jap. J. Hum. Genet., in press.

(27) Fleischer, E. A., and J. Mohr 1962 "Concerning the Genetics of the Human Haptoglobins: 126 Norwegian Families with 428 Children", Acta genet. Basel, 12, 281.

(28) Moullec, J., J. Ruffie, C. Matte, R. Audran and M. Noel. 1973 "Observations sur la Repartition des Groupes Senques (Haptoglobines, Transferrines Groups Gm) dans quelques Populations", Proc. 2nd Int. Conf. Hum. Genet., 2, 762.

(29) Sick, K.,O. Beale, D. Irvine, H. Lehmann, P. T. Goodall, and S. MacDougall. 1967 Haemoglobin $G_{Copenhagen}$ and haemoglobin $J_{Cambridge}$. Two new β-chain variants of haemoglobin A. Biochim. Biophys. Acta,140, 231.

(30) Bajatzadeh, M. and M. Walter. 1969 Studies on the Population Genetics of the Ceruloplasmin Polymorphism, Humangenetik 8, 134.

(31) Fine, J. M. 1970 Les Allotypes De L'Albumine Humaine. Etude De 8 Cas De Bis Albuminemie Observes En France, Rev. Europ. Studes Clin. et Biol. 15, 113.

(32) Rex-Kiss, B. and L. Fesus 1970 Ergebnisse Vm Bestimmungen wenschlicher Serum-transferrin-Typen, Humangenetic 10, 351.

(33) Harris, H., D. A. Hopkinson, and E. B. Robson. 1973 The incidence of rare alleles determining electrophoretic variants: data on 43 enzyme loci in man, Ann. Hum. Genet. Lond., 37, 237.

(34) Ueda, N., R. Tanis, R. Ferrell, C. Satoh, S. Kishimoto and J. V. Neel. Unpublished data.

(35) Bodmer, W. F. 1973 Population studies and the measurement of natural selection, with special reference to the HL-A system. Israel J. Med. Sci., 9, 1503.

(36) Neel, J. V. 1974 Developments in monitoring human populations for mutation rates, Mutation Res., 26, 319.

Table 1. Serum Proteins and Red Cell Enzyme Determinations in Micronesian Populations.

System	Present Investigation Tested	Present Investigation Variants	Other Investigations Tested	Other Investigations Variants	Ref.	Variants/per 1000 determinations
1. Adenosine deaminase	185	0	382	0	12	0
2. Adenylate kinase	185	0	1,387	0	12, 19	0
3. Aldolase	184	0	0	0		0
4. Carbonic anhydrase	187	0	640	4	20, 21	4.8
5. 2,3-Diphosphoglycerate mutase	185	0	0	0		0
6. Galactose-1-phosphate uridyl transferase	178	0	0	0		0
7. Indophenol "oxidase"	0	0	382	0		0
8. Isocitrate dehydrogenase	168	0	0	0		0
9. Lactate dehydrogenase	183	0	382	1	12	1.8
10. Malate dehydrogenase	186	0	382	0		0
11. Nucleoside phosphorylase	184	0	0	0		0
12. Peptidase A	185	0	382	0		0
13. Peptidase B	186	0	382	0		0
14. Phosphoglucomutase$_1$	187	0	1,387	0(49)*	12, 19	0(31.2)
15. Phosphoglucomutase$_2$	187	0	1,387	0	12, 19	0
16. 6-Phosphogluconate dehydrogenase	185	0	1,387	0	12, 19	0
17. Phosphoglycerate kinase	0	0	380	0(39)**		0(102.6)
18. Phosphohexose isomerase	185	0	382	0		0
19. Triosephosphate isomerase	182	0	0	0		0
20. Albumin	185	1	0	0		5.4
21. Ceruloplasmin	183	0	0	0	12, 13, 14, 15 16, 17, 18	0
22. Haptoglobin	185	0	2,283	0		0
23. Hemoglobin	186	0	378	0	13, 14	0
24. Transferrin	185	0	774	8	13, 15, 16	8.3
TOTAL	4,047	1	12,677	13		0.9(6.4)

*Blake, et al. (12), observed two rare types, (PGM_1^{-7} and PGM_1^{-3}) in polymorphic proportions (3-1,33; 3-2,6; 3-3,3; 7-1,33; 7-2,5; 7-3,1; 7-7,3). The PGM_1^7 type is polymorphic in several Asian populations, but the PGM_1^3 type may be a unique type among Micronesians. See discussion for elaboration.

**Blake, et al. (12), observed 39 individuals with the PGK^2 gene. This is the first report of this gene being present in polymorphic proportions in any population. See discussion for elaboration.

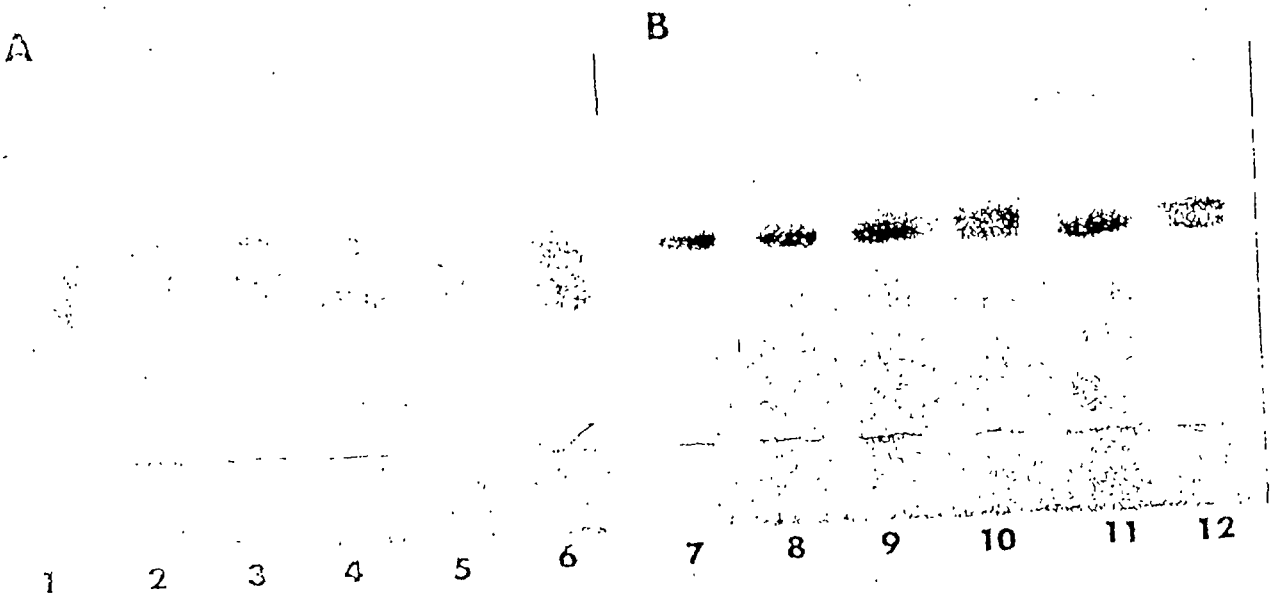
Table 2. Gene Frequencies for Six Genetic Polymorphisms in the Marshall Islands.

System	Phenotype 1	Phenotype 2-1	Phenotype 2	Total	Gene Frequency
Adenosine diaminase	167	18	0	185	$ADA^1 = 0.951$
Group specific component	116	50	2	168	$Gc^1 = 0.839$
Haptoglobin[1]	56	93	32	177	$Hp^1 = 0.579$
$Phosphoglucomutase_1^2$	156	26	1	184	$PGM_1^1 = 0.912$
Acid Phosphatase	A 101	AB 73	B 11	185	$AP^A = 0.743$
6-Phosphogluconate dehydrogenase	164	21	0	185	$6\text{-}PGD^A = 0.943$

1. The Hp^0 type was observed in 4 individuals.

2. Two examples of the PGM phenotype 2-7 and a single phenotype 1-7 were observed ($PGM_1^7 = 0.008$).

Figure 1. Albumin variant starch gel patterns using the following buffer systems: A, sodium acetate, pH 5.0; B, tris-EDTA-borate, pH 6.9. Albumin samples shown are normal serum 1, 5, 7, 8 and 11; Makiritare-2, 2; Marshall Island variant 3, 9; Makiritare-3, 4; Wapishana-1, 6, 12; Naskapi, 10. Normal samples include reference serum from our own laboratory and serum from normal Marshall Island samples. All variants, except the presently reported Marshall Island variant, have been previously compared by Tanis et al. (2).


A
B
1
2
3
4
5
6
7
8
9
10
11
12